Detection and Characterization of Chemical Vapor Fugitive Emissions from Hyperspectral Infrared Imagery by Nonlinear Optimal Estimation

Christopher M. Gittins

gittins@psicorp.com

978-689-0003

SPIE Defense, Security, and Sensing

Orlando, FL

April 5-9, 2010

Report Documentation Page					Form Approved OMB No. 0704-0188
Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.					
1. REPORT DATE **APR 2010**		2. REPORT TYPE			3. DATES COVERED **00-00-2010 to 00-00-2010**
4. TITLE AND SUBTITLE **Detection and Characterization of Chemical Vapor Fugitive Emissions from Hyperspectral Infrared Imagery by Nonlinear Optimal Estimation**					5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)					5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Physical Sciences Inc,20 New England Business Center,Andover,MA,01810-1077**					8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)					10. SPONSOR/MONITOR’S ACRONYM(S) 11. SPONSOR/MONITOR’S REPORT NUMBER(S)
12. DISTRIBUTION/AVAILABILITY STATEMENT **Approved for public release; distribution unlimited**					
13. SUPPLEMENTARY NOTES **U.S. Government or Federal Rights License**					
14. ABSTRACT					
15. SUBJECT TERMS					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**35**	

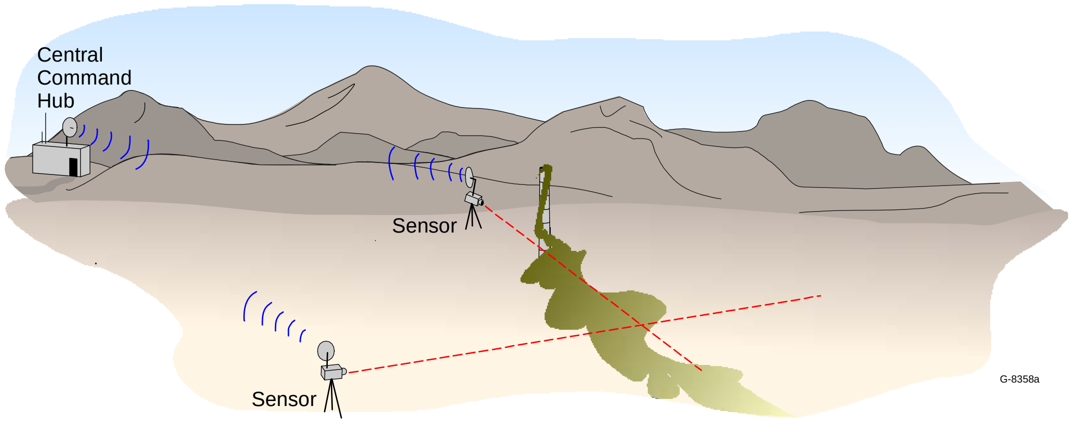
Agenda

- ***Introduction***
- **Nonlinear estimation**
  – Algorithm formulation
  – Test data
  – Results
- **Conclusions**

Algorithm Development: Overview

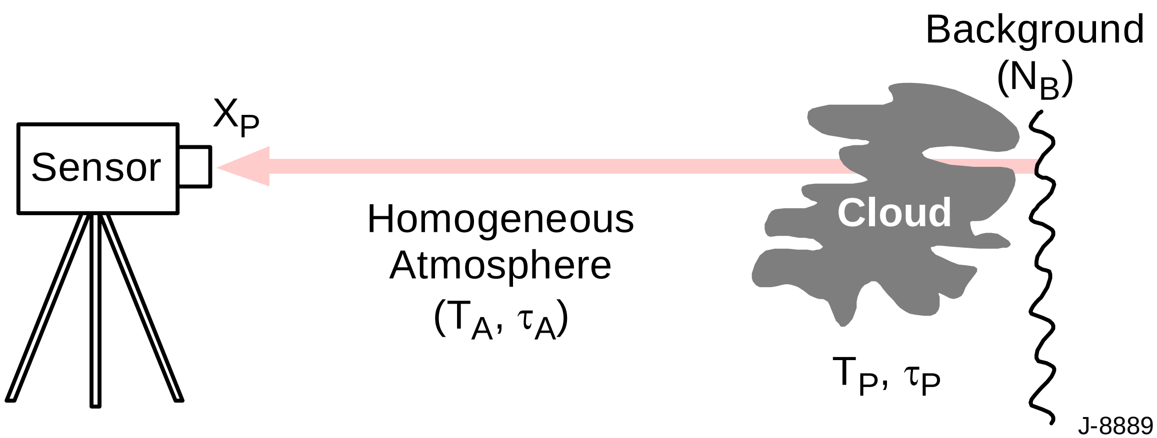
Physical Sciences Inc.

VG10-076-2



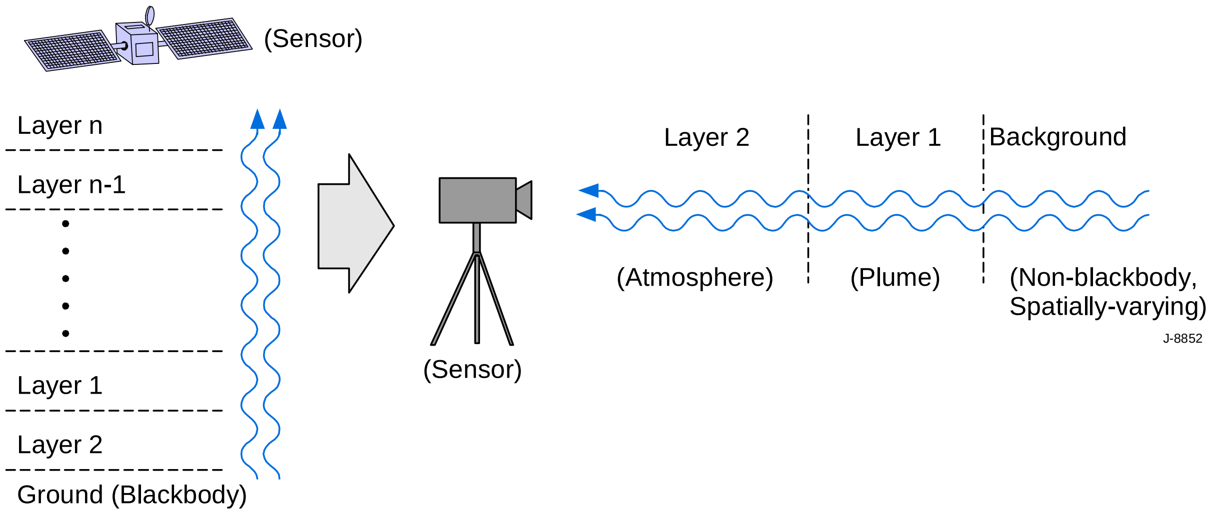
- **Objectives**
  - Improve pixel-level detection: Reduce probability of false alarm for given P_d
  - Address optically-thick plumes: Improve accuracy of estimated path integrated concentration (column density, CL)
  - Compatible with real-time processing
- **Limitations of current practice**
  - Matched-filter-based detection presumes optically-thin plume
  - Other approaches require prior measurements of background – not compatible with on-the-move detection
- ***Payoff: Improve detection immediately following large-scale release, low-lying plumes; improve mass estimate***

Problem Formulation



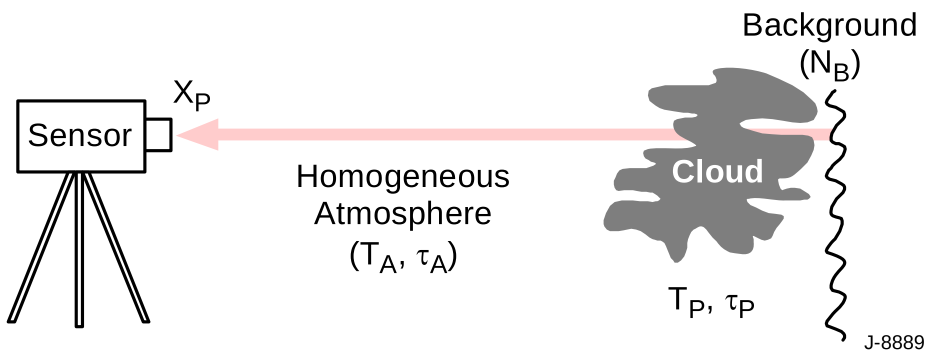
- **Ensemble of measured spectra**
- **Measured spectra are nonlinear functions of atmospheric temperature, constituent profiles, background characteristics, etc.**
- **Desire inverse solution to radiative transfer equation (RTE)**
- **Inverse solution is mathematically ill-posed – no unique solution for R_n**

Relation to Atmospheric Profile Retrieval



- **Stratified atmosphere model**
- **Profile retrieval**
  - Many stratifications
  - Simple background
  - Apply constraints to layer-to-layer variation
- **Plume detection**
  - Simple atmosphere
  - Complicated background
  - Apply constraints to background characterization

Simplified Radiative Transfer Model

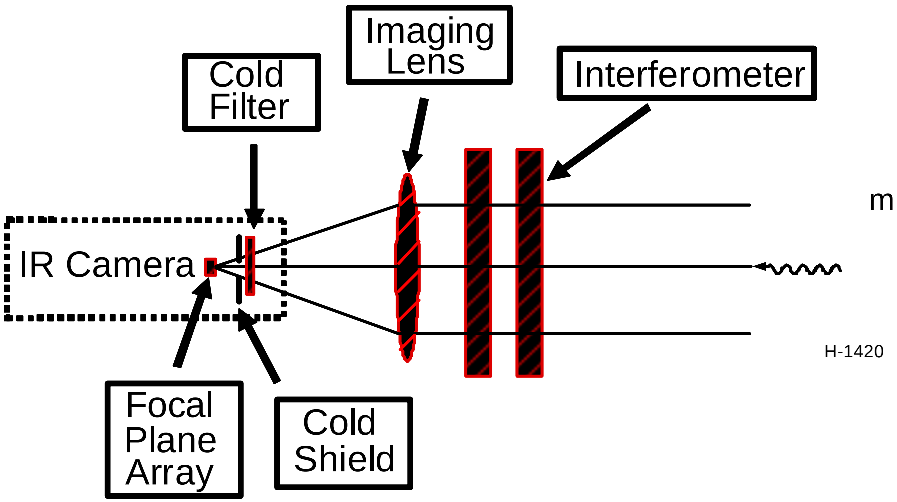


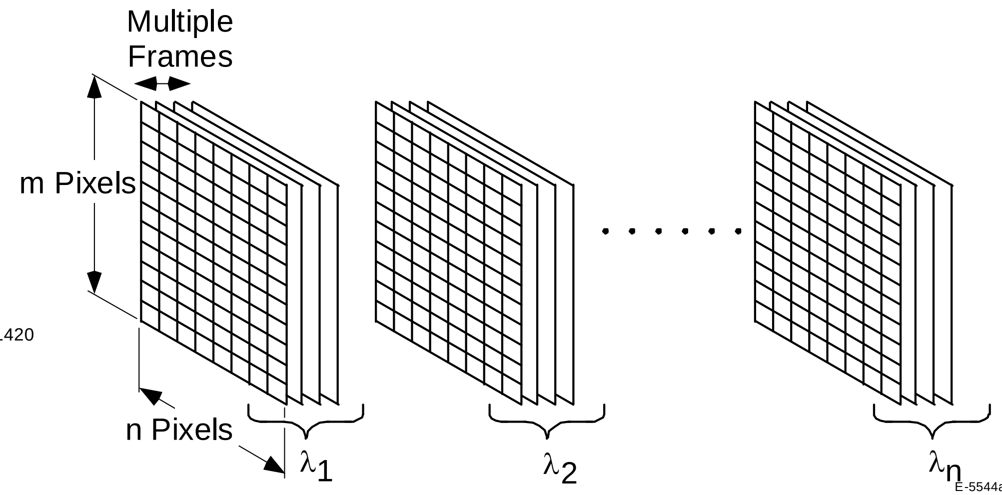
$$x_p = x_0 + \left(1-\tau_p\right)\cdot\left[\left(L_a - x_0\right) + \tau_a \cdot \left(L_p - L_a\right)\right]$$

	Linear (approx.)	Non-Linear (exact)
Plume transmission (τ_p)	**1-αs**	***exp(-αs)***
Radiance contrast (L_a-x_0)	$\propto \Delta T_{eff}$	***any***
Plume temperature (T_p)	$T_p = T_a$	$T_p = T_a$
Atmospheric scattering	**No**	**No**

- **Simplifying assumptions:**
  - Homogeneous atmosphere between sensor and vapor cloud
  - Cloud is at air temperature
- **Compare performance of non-linear (exact) RT model with linearized approximation**

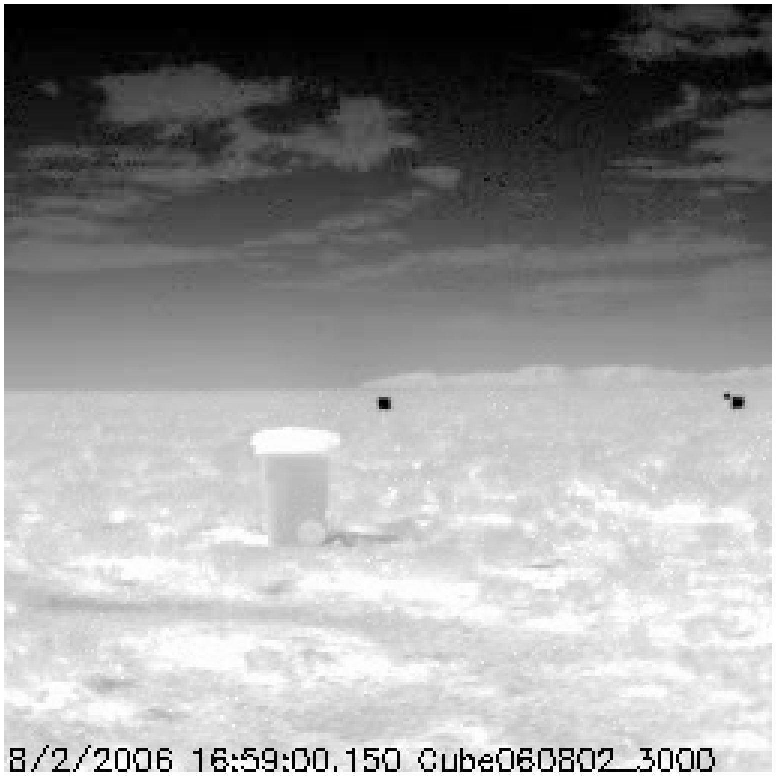
Adaptive InfraRed Imaging Spectroradiometer (AIRIS)





- **Imaging Fabry-Perot spectrometer**
  - Mirror spacing ~ λ
  - Staring IR FPA
  - Band sequential data acquisition
  - Co-registration of narrowband images
  - Tune time ~ 2 ms
- **Selective sampling of wavelengths**
  - Acquire imagery only at wavelengths which facilitate target ID
  - Minimize data volume
- **Wide field-of-view, wide spectral coverage**

TEP Detection:
Shortcoming of Thin Plume Approximation



- **Triethyl phosphate (TEP) release**
- **Post-processing:**
  - Non-linear estimator in IDL
  - False alarm mitigation: 4 of 8 spatial filter
  - Bad pixels substituted
- **Detection key:**
  - TEP only
  - Yellow: OD ~ 0
  - Red: OD $\geq$ 1

Agenda

- **Introduction**
- ***Nonlinear estimation***
  - *Algorithm formulation*
  - Test data
  - Results
- **Conclusions**

Optimal Estimation: Bayesian Approach

- **Bayesian posterior pdf for model parameter values:**

$$\boxed{p(\theta|\tilde{x}) = \frac{p(\tilde{x}|\theta)p(\theta)}{p(\tilde{x})}}$$

- **Maximum likelihood parameter values maximize posterior:**

$$\hat{\theta} = \arg\max\{p(\theta|\tilde{x})\}$$
$$= \arg\min\{-\ln p(\theta|\tilde{x})\}$$

- **Multi-variate normal pdf for deviation between model and measurement:**

$$-\ln p(\tilde{x}|\theta) = \tfrac{1}{2}[\tilde{x} - f(\theta)]^T D^{-1}[\tilde{x} - f(\theta)] + c_{x|\theta}$$

$$D = diag\{[\sigma_1^2, \sigma_2^2, \ldots, \sigma_k^2]\}$$

- **Prior pdf for model parameter values**

$$-\ln p(\theta) = \tfrac{1}{2}[\theta - \theta_a]^T R[\theta - \theta_a] + c_\theta$$

Optimal Estimation: Signal Model

- **Signal model:**

$$f(\theta) = \tau_e \circ x_0 + [1 - \tau_e] \circ L_a$$

calculated spectrum | calculated plume trans. | estimated “baseline” | Blackbody at T_{atm}

- **Plume transmission:**

$$\tau_e = \exp[-\alpha s]$$

peak OD | reference spectrum

- **Infrared background:**
  - Linear mixing model
  - Probabilistic Principal Components
  - Robust estimate of sample covariance (Huber-type M-estimator)

$$\boxed{x_0 = \mu + B\beta}$$

- **Model parameters:**
  - α: Plume OD
  - T_a: Plume/air temperature
  - β: Parameters which account for bkgd. radiance given bkgd. model

$$\boxed{\theta = [\alpha, T_a, \beta]}$$

Minimize Cost Function

- **Maximum likelihood parameter values minimize cost function**
- **Multivariate normal pdfs result in "quadratic" cost function**

$$C = \left[\tilde{x} - f(\theta)\right]^T D^{-1}\left[\tilde{x} - f(\theta)\right] + \left[\theta - \theta_a\right]^T R_\theta \left[\theta - \theta_a\right]$$

deviation between measured and model spectra

deviation of parameters from nominal values

  - Quadratic formulation: $C = r^T r$
  - Prior applied to background coefficients only: $\boxed{\left[\theta - \theta_a\right]^T R_\theta \left[\theta - \theta_a\right] = \beta^T \beta}$
  - Residuals vector: $r \equiv \left[D^{-1/2}\left[\tilde{x} - f(\theta)\right]; \beta\right]$

- **Determine maximum likelihood parameter values by nonlinear estimation**
  - Approach not limited to quadratic cost function
  - Quadratic cost function amenable to computationally-efficient solution

Nonlinear Optimization Algorithms

- **Iterative determination of parameters, e.g., Newton's Method:**

cost function

$$\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \mathbf{H}_i^{-1} \left(\nabla C \right)_i$$

Model parameters

iteration no. Hessian matrix gradient vector

- **Gauss-Newton algorithm**
  - Appoximate Hessian matrix: $\mathbf{H} \approx 2\mathbf{J}^T \mathbf{J}$
  - Parameter update equation: $\boxed{\boldsymbol{\theta}_{i+1} = \boldsymbol{\theta}_i - \left(\mathbf{J}_i^T \mathbf{J}_i \right)^{-1} \mathbf{J}_i^T \mathbf{r}_i}$
  - Initial guess at θ from linear model
- **Levenberg-Marquardt algorithm also applicable**

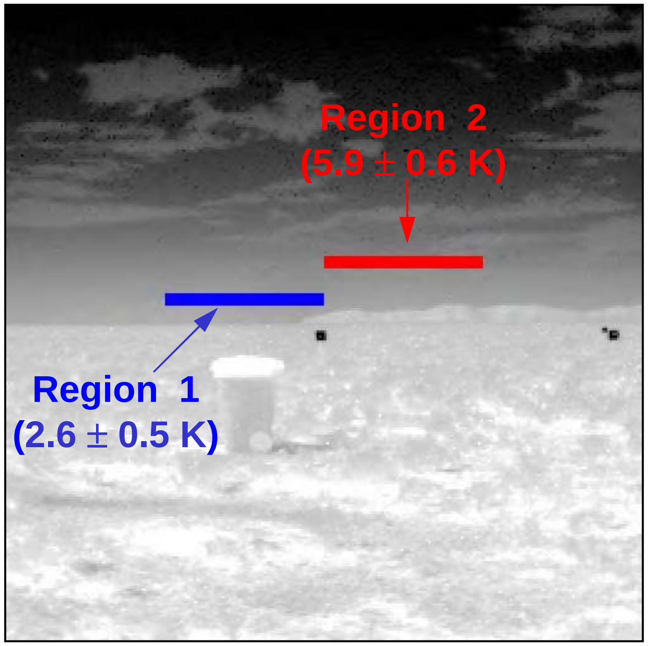
Agenda

- **Introduction**
- ***Nonlinear estimation***
  - Algorithm formulation
  - ***Test data***
  - Results
- **Conclusions**
- **Next generation algorithm(s)**

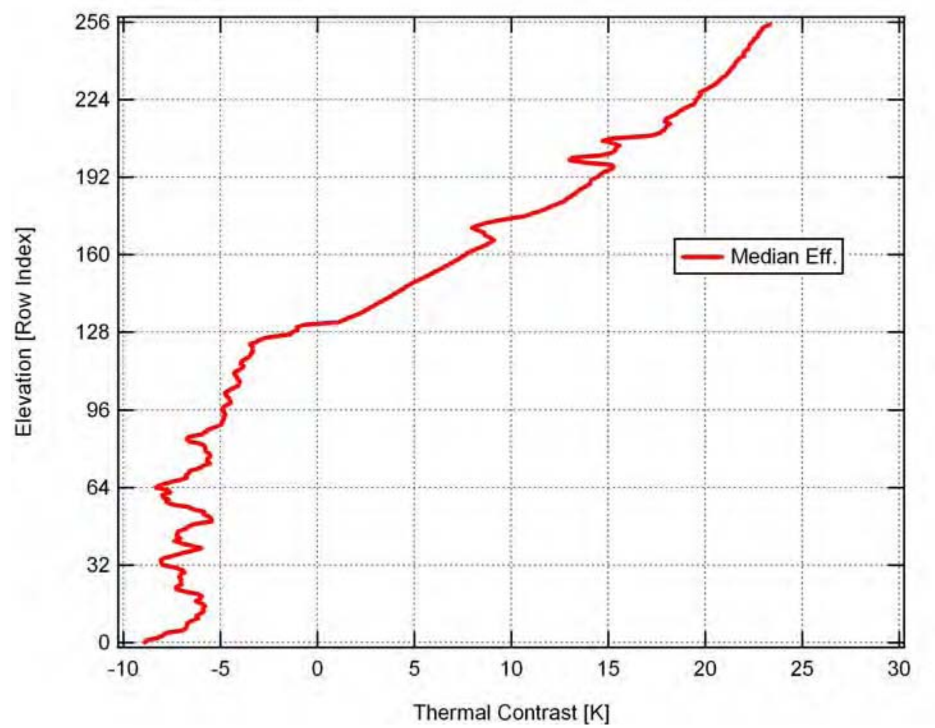
Test Regions

Physical Sciences Inc.

VG10-076-14



AIRIS-WAD datacube: 256 x 256 pixels



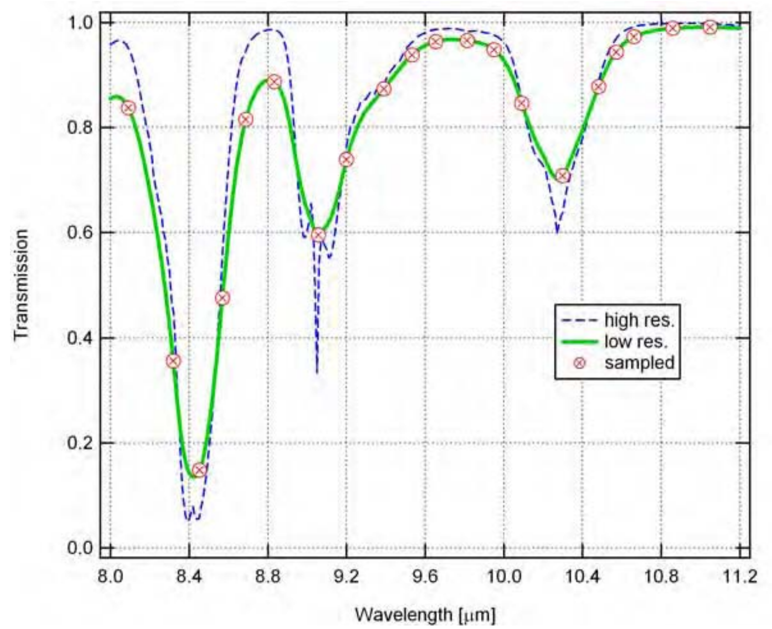
- **Plume-free data augmented with synthetic plumes:**
  - 64 x 5 pixels
  - Max OD from 0 to 3.0 (base e)
  - T(plume) = T(air) = 25.0 deg C
- **Thermal contrast**
  - ~0 K along horizon
  - Monotonic increase with elev. angle
- **Test both favorable and unfavorable regions**

Simulation: Synthetic R-134a Plumes

- **Effective plume transmission:**
  - Reference spectrum from PNNL library

$$\tau(\lambda) = \exp[-CL \cdot \sigma(\lambda)]$$

  - Specify column density
  - Beer's Law + instrument resolution function



$$\bar{\tau}_p(\lambda_s) = \int \tau(\lambda) \cdot g(\lambda, \lambda_s) \cdot d\lambda$$

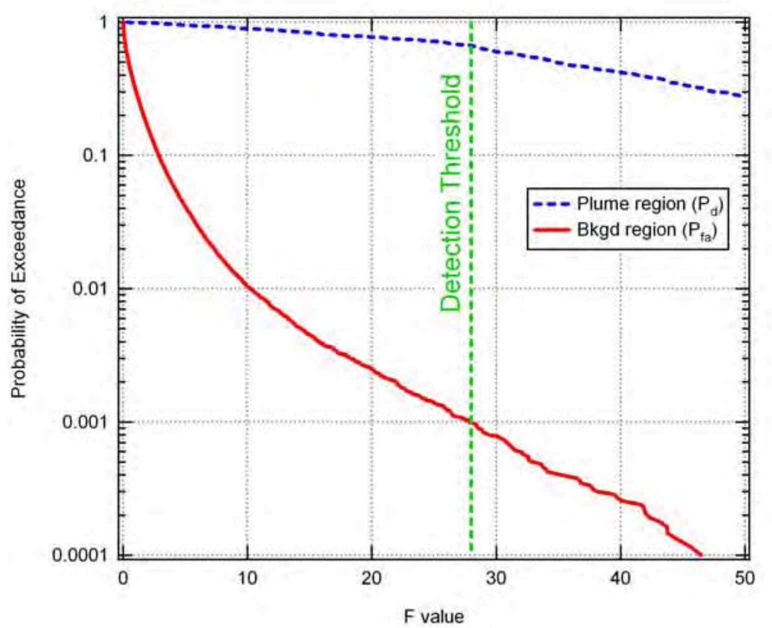
- **Data augmentation:**
  - Partition measurement into estimated signal, noise
  - Modify signal w/plume signature
  - Add back estimated noise

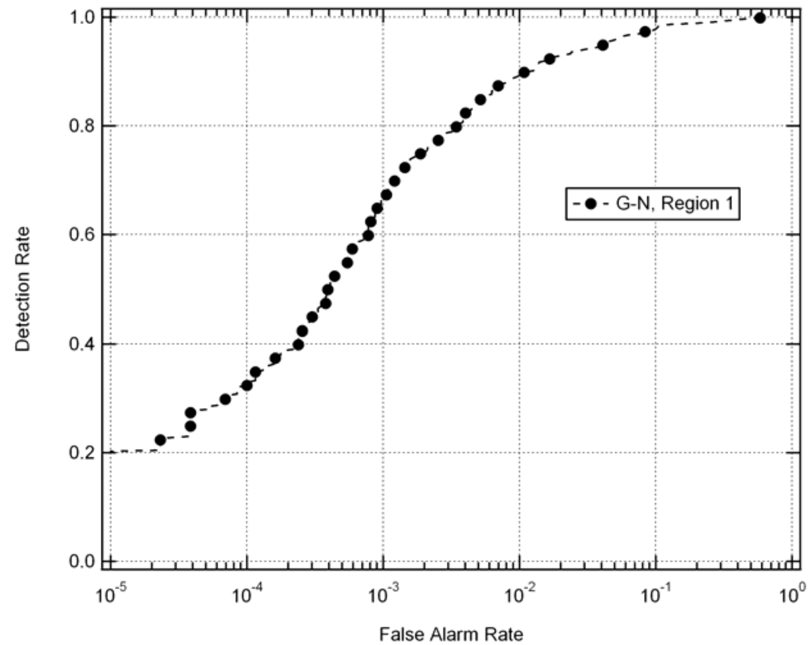
$$x_p = \hat{x}_0 + [1 - \tau_p] \circ [L_a - \hat{x}_0] + \hat{e}$$

Performance Metric: ROC Curves

Physical Sciences Inc.

VG10-076-16





- **Binary decision hypotheses**
  - H_0 ("plume absent") and H_1 ("plume present")
  - pdfs for detection statistic: $p(F \mid H_i)$
- **ROC curve is $P_d(F_{th})$ vs $P_{fa}(F_{th})$**
  - P_d from plume-augmented region
  - P_{fa} from rest of scene
- **ROC "surface": $P_d(\alpha; F_{th})$**

$$P_{fa}(F_{th}) = \int_{F_{th}}^{\infty} p(F \mid H_0) d\eta$$

$$P_d(F_{th}) = \int_{F_{th}}^{\infty} p(F \mid H_1) d\eta$$

Agenda

- **Introduction**
- ***Nonlinear estimation***
  - Algorithm formulation
  - Test data
  - ***Results***
- **Conclusions**

Performance Comparison with Matched Filter

Physical Sciences Inc.

- **Objective: Compare nonlinear estimation with matched filter estimation**
  - Detection statistics
  - Column density/optical density
- **Detection with nonlinear estimator: F test**

$$F(\tilde{x}) = (k-1)\cdot\left[\frac{C(\tilde{x},\hat{\theta}_0)}{C(\tilde{x},\hat{\theta})} - 1\right]$$

- **Analogous metric for clutter-matched filter: Adaptive Cosine Estimator (ACE)**

$$D_{MF}(\tilde{x}) = \frac{\left(s'^T\,\hat{\Sigma}^{-1}[\tilde{x}-\mu]\right)^2}{\left(s'^T\,\hat{\Sigma}^{-1}s'\right)\left([\tilde{x}-\mu]^T\,\hat{\Sigma}^{-1}[\tilde{x}-\mu]\right)}$$

$$F_{MF} = (k-1)\frac{D_{MF}}{1-D_{MF}}$$

- **Matched-filter optical density estimate:**

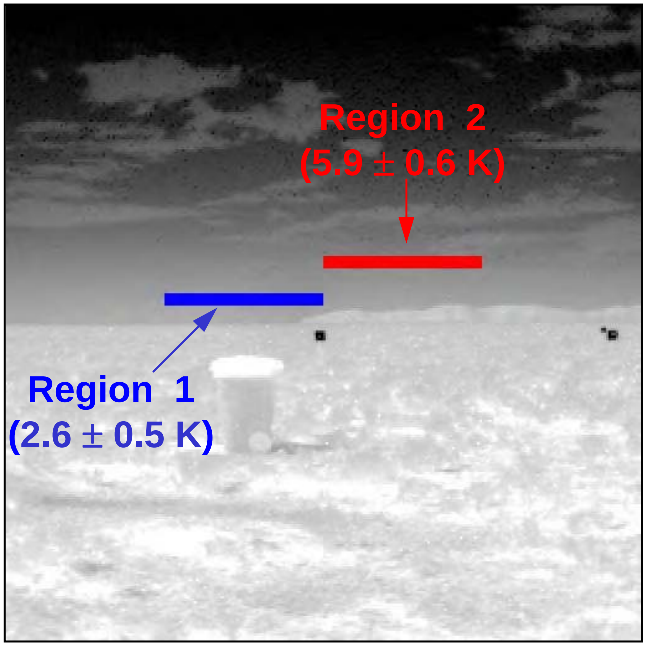
$$\hat{\alpha}_{MF} = \frac{s'^T\,\hat{\Sigma}^{-1}(\tilde{x}-\mu)}{s'^T\,\hat{\Sigma}^{-1}s'}\cdot\frac{\Delta T_0}{\Delta T_{eff}}$$

- ***Expect near identical results for optically-thin plumes***

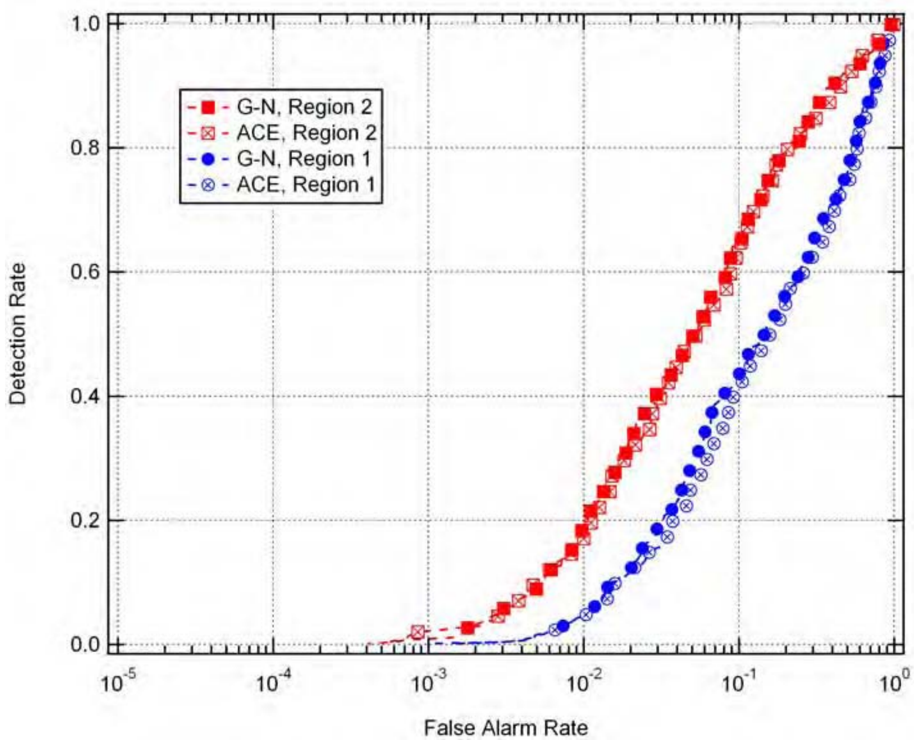
R-134a Detection: Optically-Thin Plume, OD=0.1

Physical Sciences Inc.

VG10-076-19



AIRIS-WAD datacube: 256 x 256 pixels

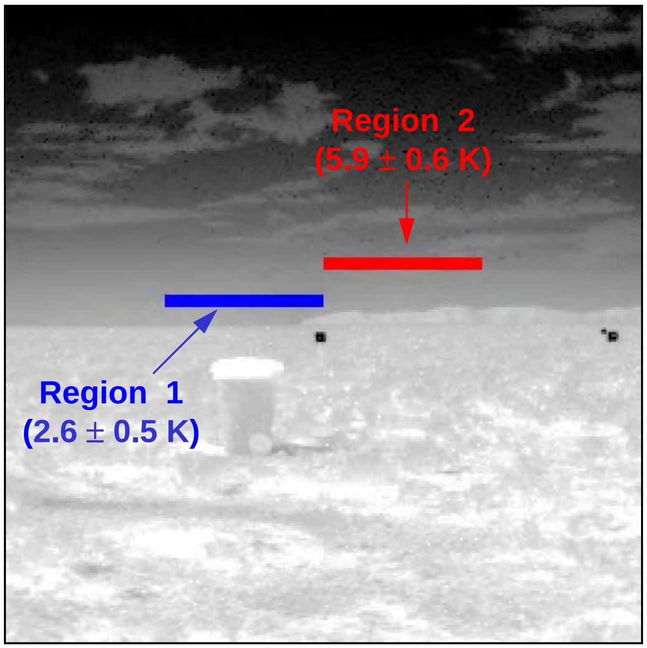


- **Plume column density = 82 mg/m² (20 ppmv-m)**
- **Detection statistics not favorable in either Region**
- **ACE and Gauss-Newton ROC curves are nearly identical**
  - 20 bands in test datacube
  - OD=0 reference spectrum

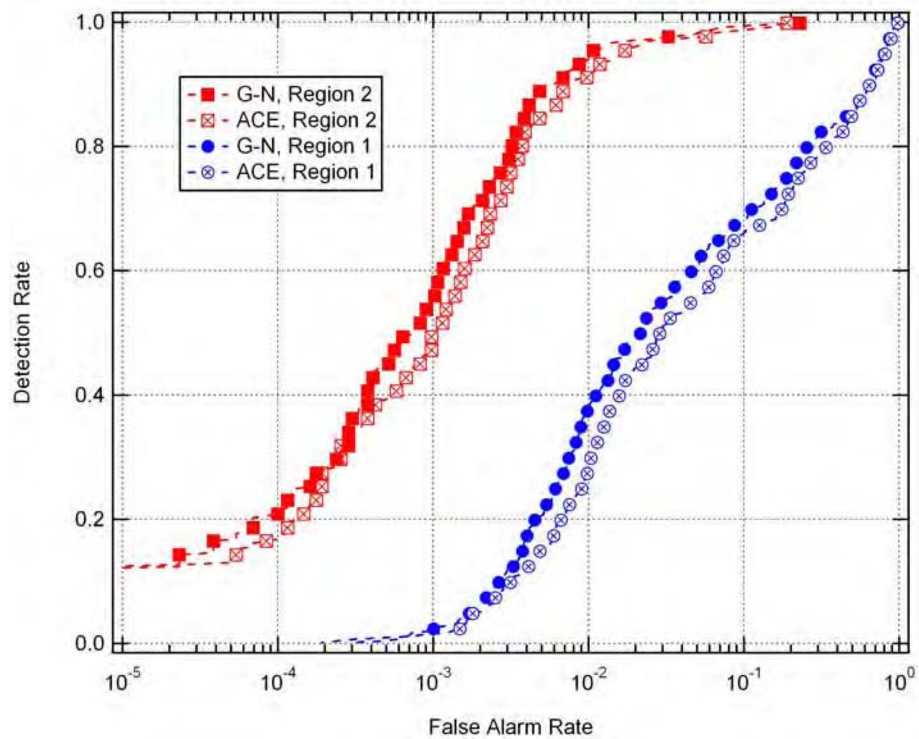
R-134a Detection: Optically-Thin Plume, OD=0.3

Physical Sciences Inc.

VG10-076-20

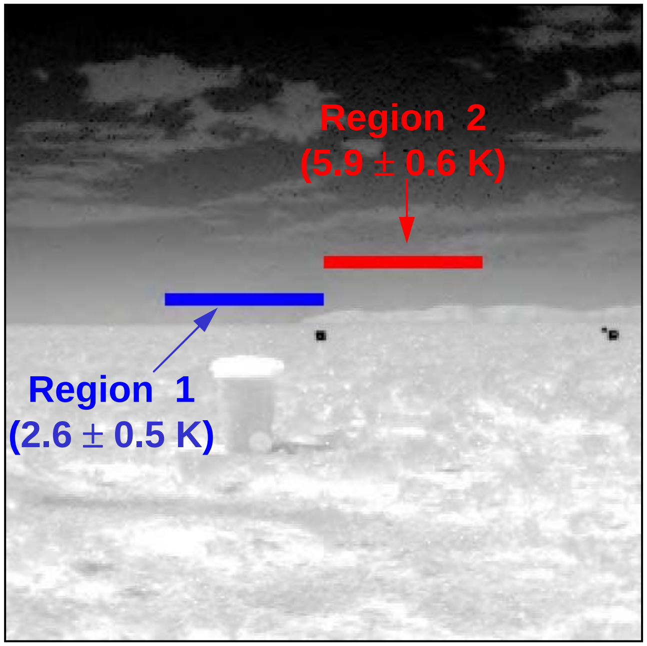


AIRIS-WAD datacube: 256 x 256 pixels

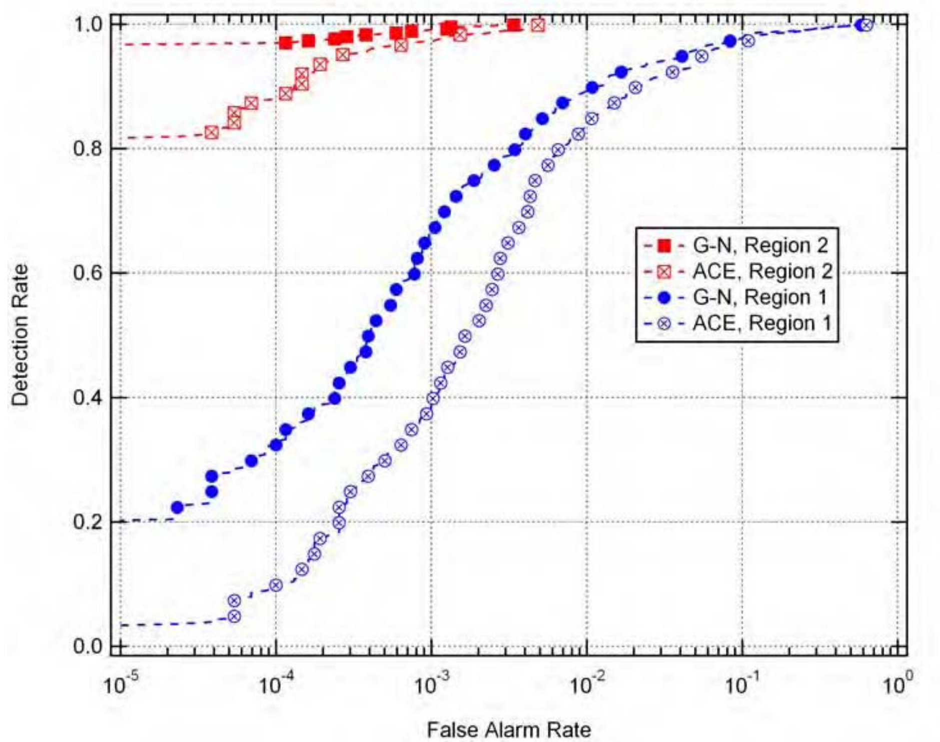


- **Plume column density = 246 mg/m² (59 ppmv-m)**
- **Detection statistics not favorable in Region 1, marginal in Region 2**
  - Lower thermal contrast
  - ~2 orders of magnitude reduction in P_{fa} from Region 1 to Region 2
- **ACE and Gauss-Newton ROC curves are nearly identical**

R-134a Detection: Optically-Thick Plume, OD=1.0



AIRIS-WAD datacube: 256 x 256 pixels

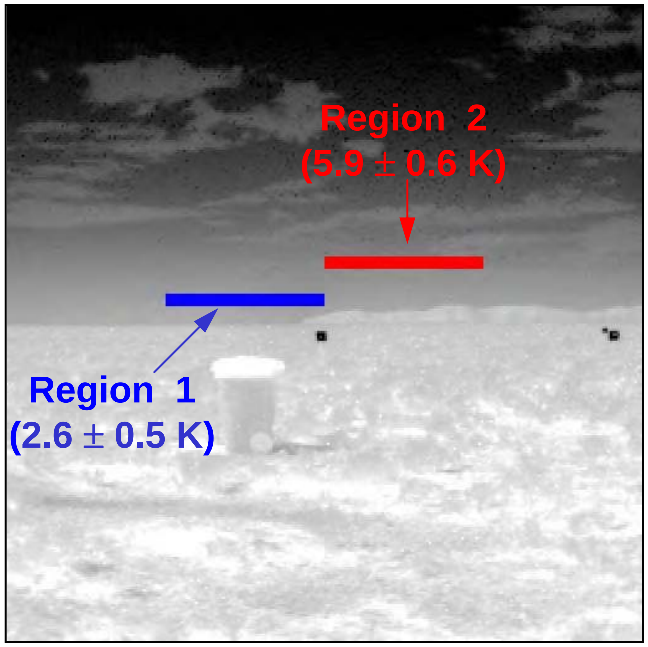


- **Plume column density = 822 mg/m² (197 ppmv-m)**
- **Detection statistics favorable in Region 2, marginal in Region 1**
  - >2 orders of magnitude reduction in P_{fa} from Region 1 to Region 2
- **Gauss-Newton produces significantly more favorable ROC curves than ACE**
  - Factor of ~2 improvement in Region 1 (P_{fa} for fixed P_d)
  - Multiple orders of magnitude improvement in Region 2

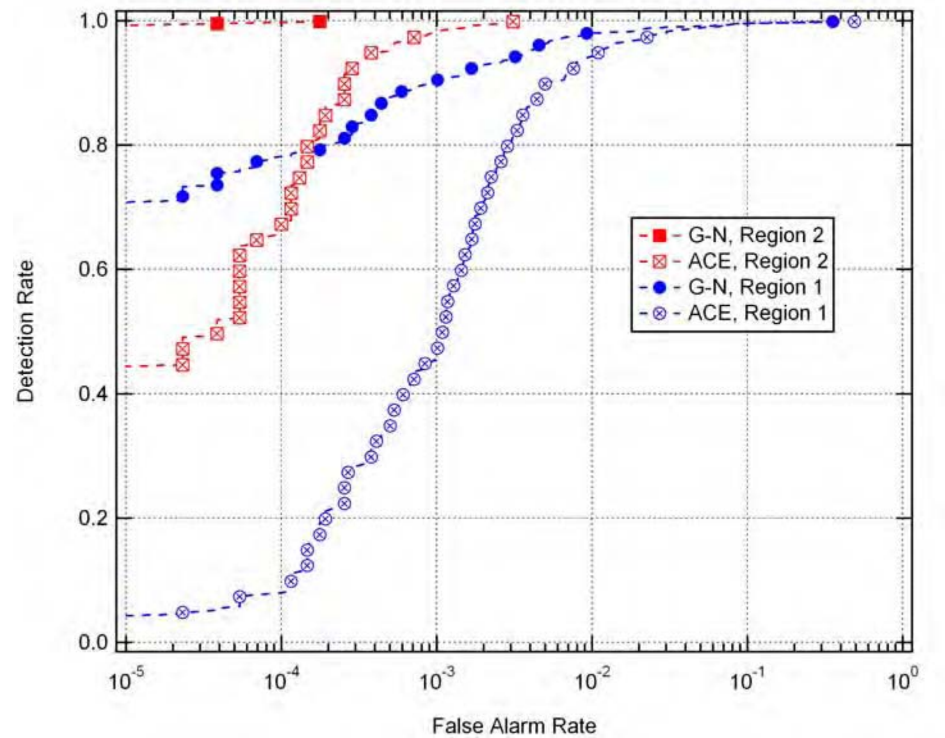
R-134a Detection: Optically-Thick Plume, OD=2.0

Physical Sciences Inc.

VG10-076-22



AIRIS-WAD datacube: 256 x 256 pixels

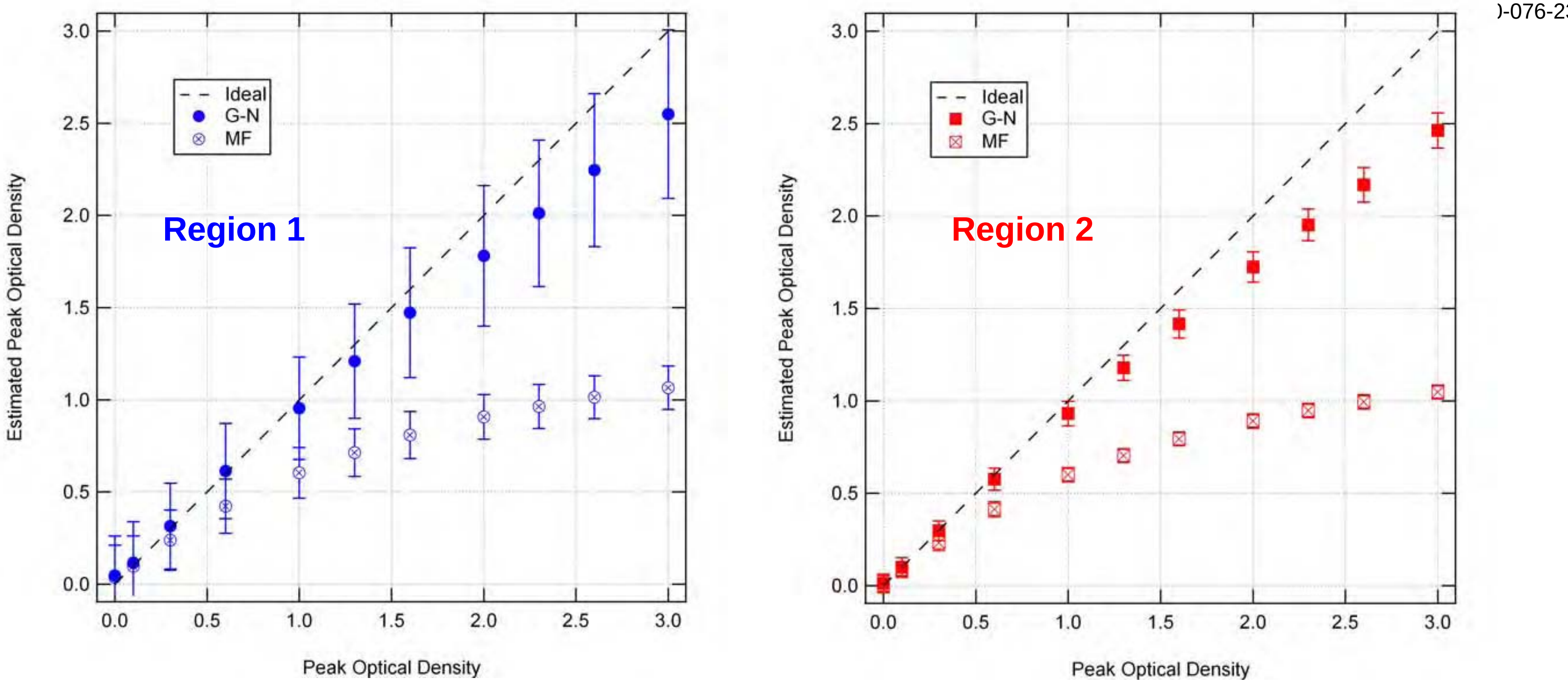


- **Plume column density = 1643 mg/m² (394 ppmv-m)**
- **Detection statistics favorable in both Regions**
- **Gauss-Newton produces significantly more favorable ROC curves than ACE**
  - >1 order of magnitude improvement in Region 1
  - Multiple orders of magnitude improvement in Region 2

Column Density Estimation

Physical Sciences Inc.

)-076-23



- **Increased thermal contrast reduces uncertainty, no effect overall accuracy**
- **Nonlinear estimation**
  - Accurately recovers embedded OD (CL)
  - Systematic deviation at OD>1 is instrument resolution effect
- **Matched Filter systematically underestimates CL**
- ***Nonlinear estimator always as good or better than MF***

Algorithm Execution

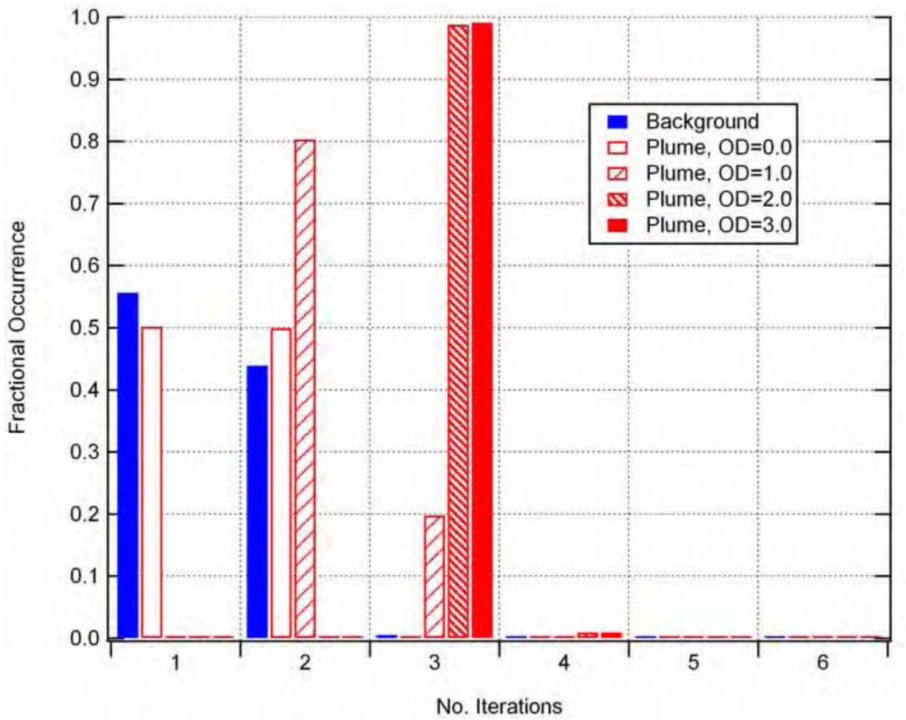
Physical Sciences Inc.

VG10-076-24

- **Gauss-Newton algorithm is iterative**
- **Termination criterion:**

$$0 < \left[1 - \frac{C_{i+1}}{C_i}\right] < \delta_{\max}$$

- **Initial guess is Iteration 0**
- **Typical results:**
  – 1-2 iterations for no plume (plume OD=0)
  – 3 iterations to converge for OD~2-3 TEP plume
- **Decreasing δ to 0.0001 increase no. iteration but no statistically-significant effect on CL**



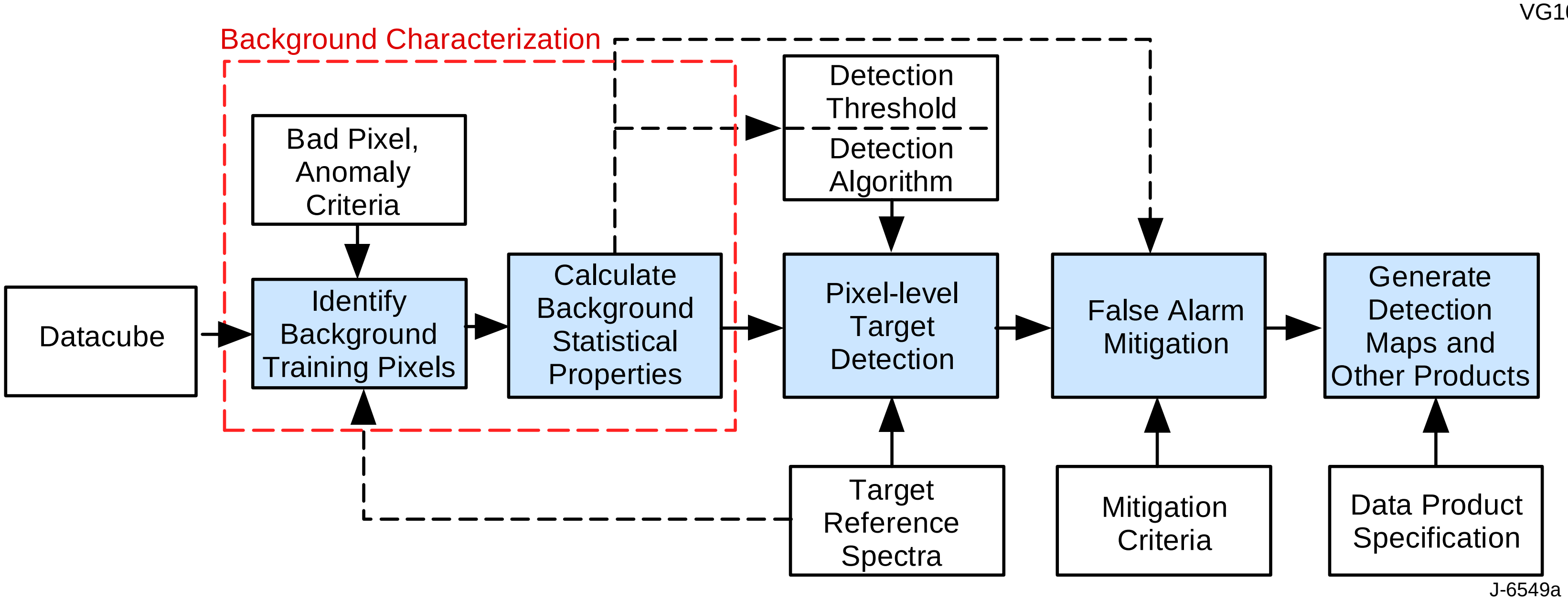
δ_{max} = 0.01

Summary and Conclusions

- **Developed nonlinear estimator for plume detection and characterization based on RTE**
  - Bayesian formulation
  - Statistical model for IR background
  - Gauss-Newton algorithm to estimate maximum *a posteriori* (MAP) values
- **Signal model developed for non-scattering atmosphere, single layer plume**
  - Easily modified to address more complicated atmospheres
- **Nonlinear estimation significantly outperforms matched-filter-based with optically-thick plumes**
  - "Orders of magnitude" improvement
  - NL estimator and matched filter produce equivalent results for optically-thin plumes
- **This work was performed under Contracts from the Defense Threat Reduction Agency (HDTRA01-07-C-0067) and US Army ECBC Aberdeen Proving Ground, MD (W911SR-06-C-0022). Any opinions, findings and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of HDRA or the Army.**

Additional Material

Data Processing Chain



- **Focus is pixel-level target detection**
- **New background characterization approach facilitates improved pixel-level detection**
- **"A chain is only as strong as its weakest link."**
  - Provide higher quality input to False Alarm Mitigation block
  - False Alarm Mitigation is separate issue

Technical Approach

- **Adapt methodology used for atmospheric profile retrieval from space-based sensor data (e.g. AIRS, IASI, MODIS, TES)**
  - Parameterize Radiative Transfer Equation (RTE)
  - Apply Estimation Theory to determine max. likelihood parameter values
  - Exploit large data set: utilize ensemble statistics
- **Rationale:**
  - Physics-based model for observations
  - Statistically-justified constraints
  - Strong theoretical foundation (see, e.g., C.D.Rodgers, <u>Inverse Methods for Atmospheric Sounding</u>)
- **Benefits**
  - Adaptable framework
  - Immediate application to non-scattering atmosphere
  - Can modify RTE to address more complicated atmospheres

Linear Models

$$x_p = x_0 + \alpha' s'$$

$$x_0 = \mu + B\beta \qquad\qquad x_0 = \nu_b \sim N(\mu, \Sigma)$$

"Structured background" "Unstructured background"

- **"Structured Background"**
  - Values of β are unconstrained
  - Generalized Likelihood Ratio Test: $D_{GLRT}(x) = \dfrac{x^T P_B^{\perp} x}{x^T P_{SB}^{\perp} x}$ $\qquad P_B^{\perp} = I - B\left(B^T B\right)^{-1} B^T$
  - Typical implementation: B = eigenvectors of sample covariance matrix
- **"Unstructured Background"**
  - ν_b is a random vector
  - Adaptive Cosine Estimator: $D_{ACE}(x) = \dfrac{\left[s'^T \Sigma^{-1} x\right]^2}{\left[s'^T \Sigma^{-1} s'\right]\left[x^T \Sigma^{-1} x\right]}$
- **Survey article: Manolakis, Marden, & Shaw, "Hyperspectral Image Processing for ATR Applications," *Lincoln Lab J.*, v.14 (2003)**

Pros and Cons of Linear Approximation

Physical Sciences Inc.

VG10-076-30

- **Pro: Matrix multiplication results in fast computation**
  - All spectra in ensemble may be processed in parallel
  - Major computational expense is diagonalization of sample covariance matrix
  - *AIRIS-WAD: <150 ms to process 65536 twenty element spectra for four target signatures (using 2005 vintage technology)*

- **Pro: Detection statistics well-understood for Gaussian noise**

- **Con: Underlying physical assumptions not valid for detection scenarios of interest**
  - Mathematical model not matched to physics

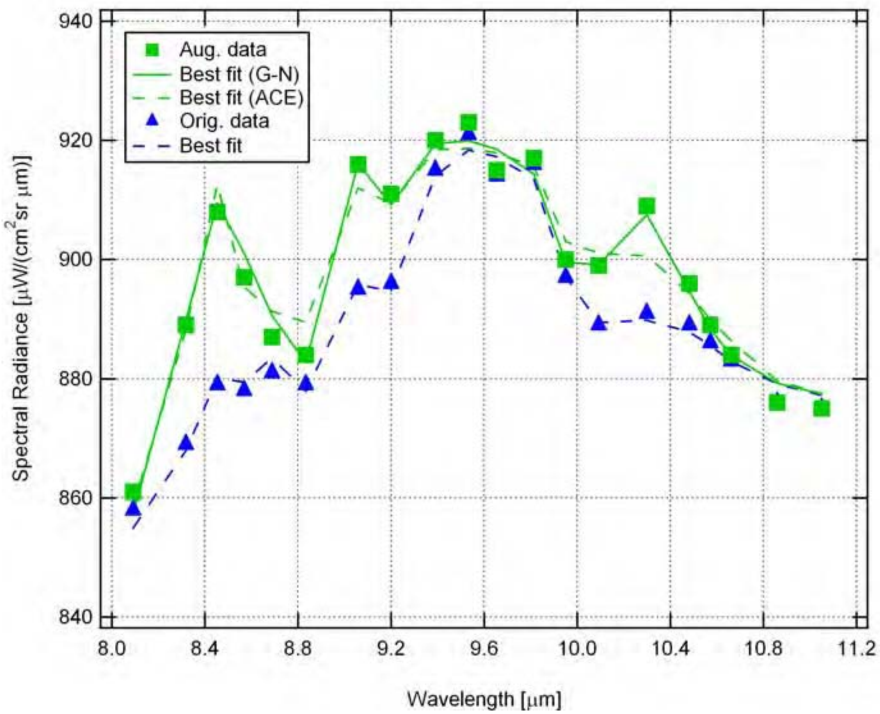
$$\tau_p = \exp(-\alpha s) \approx \cancel{1 - \alpha s}$$

  - Linear approximation to Beer's Law can introduce significant error

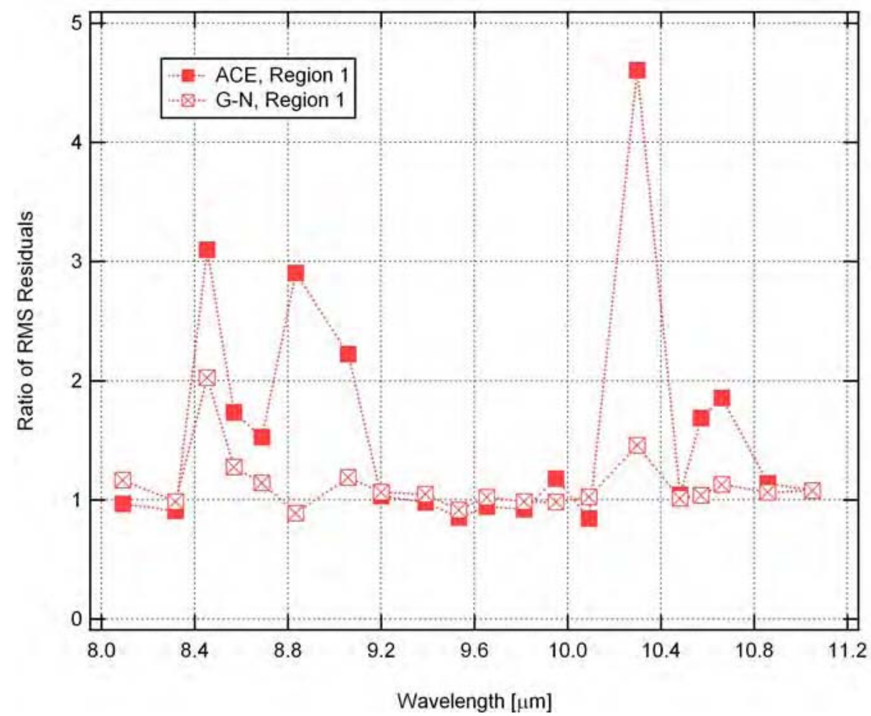
Why Gauss-Newton Yields Better Results

Physical Sciences Inc.

VG10-076-31



Spectrum augmented with OD=3.0 plume



Ratios of rms residuals in plume region, OD=3.0 to OD=0

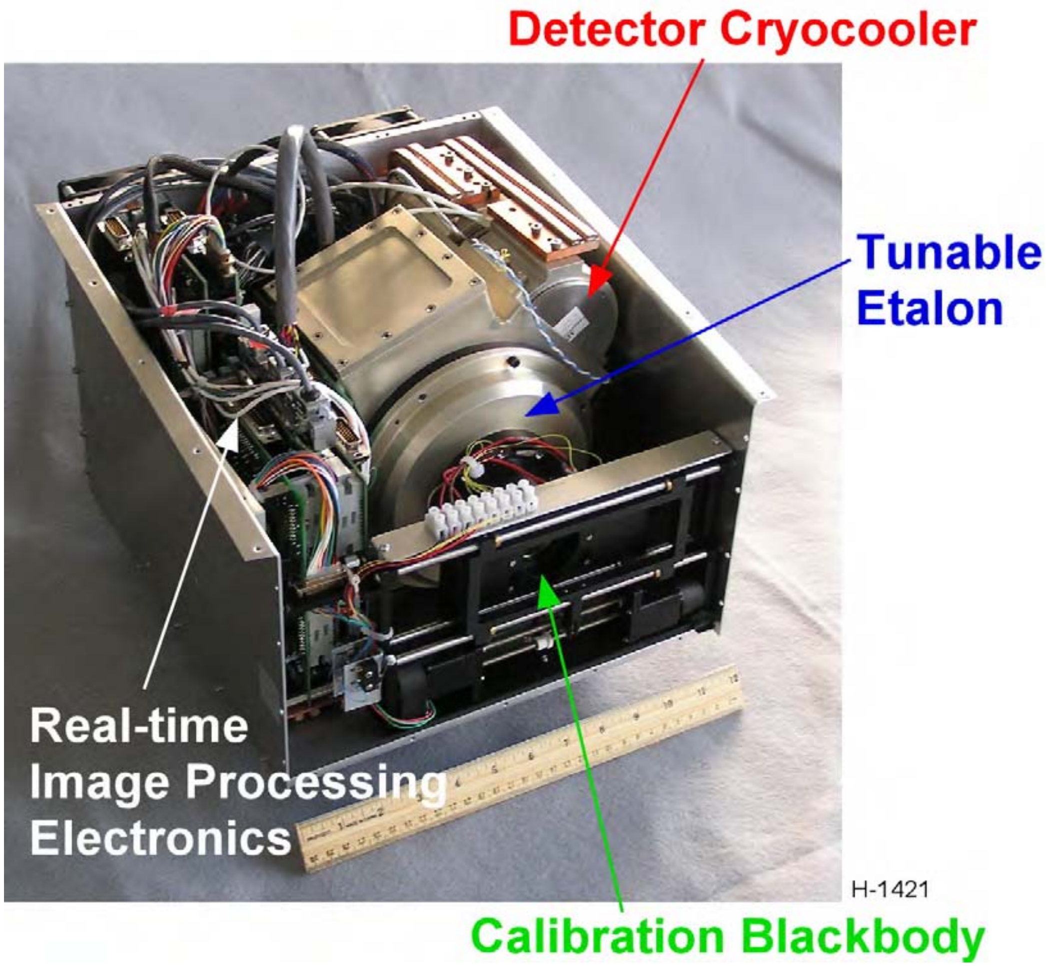
- **Model is matched to the data**
- **Fit residuals are systematically larger with linear model**
  - Result of least-squares minimization
  - Location of largest residuals highly correlated with strongest R-134a absorption features

Adaptive Infrared Imaging Spectroradiometer – Wide Area Detector (AIRIS-WAD)

Physical Sciences Inc.

VG10-076-32

- **Optical:**
  - 256 x 256 pixels
  - 30 deg x 30 deg FOV
  - spectral coverage: 7.9 to 11.2 μm at ~0.1 μm resolution (~1% of λ)
- **Datacubes:**
  - 20 wavelengths
  - user selectable λ's, specified prior to mission
- ***Real-time datacube processing: up to 3 Hz***
- **Detection algorithm history:**
  - GLRT: Winter 2005-Spring 2006
  - ACE: since Spring 2006



Hyperspectral Background Model

Physical Sciences Inc.

VG10-076-33

- **Probabilistic Principal Components-based**
  - M.E.Tipping & C.M.Bishop, *J.R.Statist. Soc. B* (1999)
- **Linear mixing model**

$$x = \mu + B\beta$$

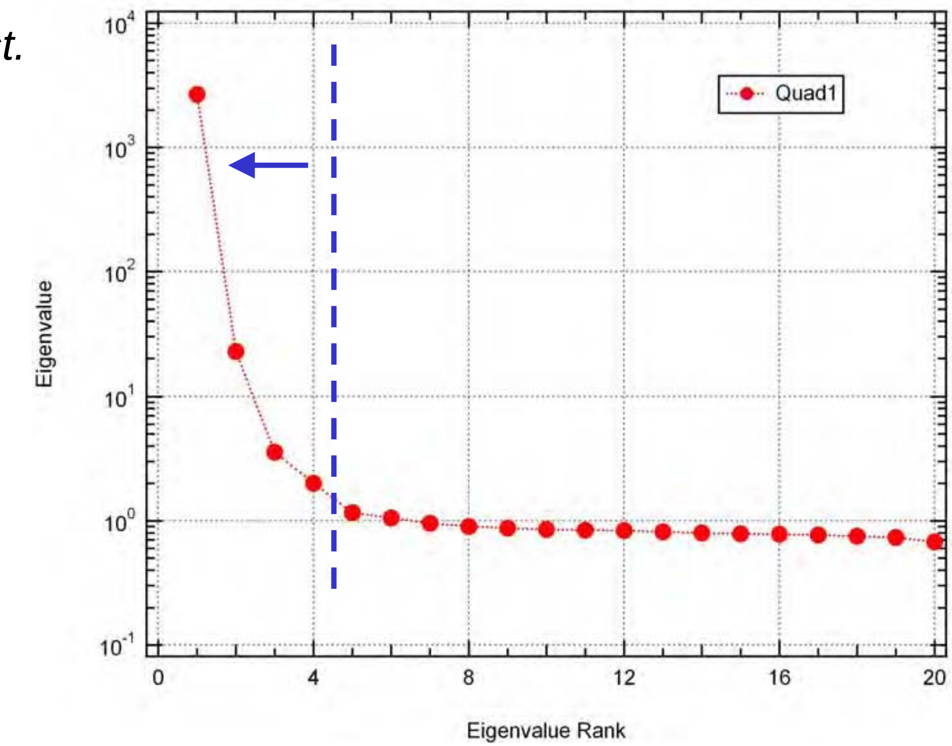
- **Eigenvalue-based covariance regularization**

$$\Sigma \approx \hat{\Sigma} = BB^T + \varepsilon D$$

$$\Sigma = D^{1/2}\left(U\Lambda U^T\right)D^{1/2}$$

$$B = D^{1/2}U_m\left(\Lambda_m - \varepsilon I_m\right)^{1/2}$$

- ***Σ = robust estimate of sample covariance: Huber-type M-estimator***



Gauss-Newton Algorithm

- **Follows from Newton's method – simplifying approximations**
- **Good for solving weakly nonlinear equations**
- **Hessian matrix:**

$$H_{jk} = 2\sum_{q=1}^{m}\left[\frac{\partial r_q}{\partial \theta_j}\frac{\partial r_q}{\partial \theta_k} + r_q \frac{\partial^2 r_q}{\partial \theta_j \partial \theta_k}\right]$$

$$\approx 2\sum_{q=1}^{m}\left[\frac{\partial r_q}{\partial \theta_j}\frac{\partial r_q}{\partial \theta_k}\right] = 2J^T J$$

$$J = \frac{\partial r}{\partial \theta}$$

Jacobian

- **Gradient:** $$[\nabla_\theta C]_j = \frac{\partial C}{\partial \theta_j} = 2\sum_{q=1}^{m}\left[r_q \frac{\partial r_q}{\partial \theta_j}\right]$$

$$\nabla_\theta C = 2J^T r$$

- **Parameter update equation:** $$\boxed{\theta_{i+1} = \theta_i - \left(J_i^T J_i\right)^{-1} J_i^T r_i}$$

- **Initial guess at θ from linear model**